Conventional solution and proof

For each subset A, let v_A be its indicator vector indexed by the finite universe and let G be the base Gram matrix. Then $K'(A,B) = v_A^\top G v_B$. For sets A_i and coefficients c_i,

$$\sum_{i,j} c_i c_j K'(A_i, A_j) = \Big(\sum_i c_i v_{A_i}\Big)^\top G \Big(\sum_j c_j v_{A_j}\Big) \geq 0.$$

The kernel is symmetric because G is symmetric. Empty sets have zero feature and cause no exception.

Structural or analytic derivation

A base feature map gives the aggregate feature $\Phi(A) = \sum_{x \in A} \phi(x)$. The set kernel is $\langle \Phi(A), \Phi(B) \rangle$. This proves that the construction sums features without requiring pairwise expansion on each repeated query.

Comparative cost and preferred method

Direct evaluation uses $|A||B|$ base-kernel calls. With an available d-dimensional base feature, aggregate construction costs $O(d(|A|+|B|))$ and each subsequent comparison costs $O(d)$. The advantage requires feature access and enough reuse; an arbitrary black-box kernel need not provide a small finite feature map.